

Award No. G11AP20163

National Geological and Geophysical Data Preservation Program

**Washington State Metadata Project:
Washington Metal Mines Metadata Production**

Final Technical Report

Robert Berwick/Stefanie Kinnamon

Washington Department of Natural Resources

Division of Geology and Earth Resources

PO Box 47007

Olympia, WA 98504-7007

(360) 902-1431 / (360) 902-1469

(360) 902-1785 fax

robert.berwick@dnr.wa.gov

stefanie.kinnamon@dnr.wa.gov

Award Term: July 1, 2011 to June 30, 2012

Submitted: September 11, 2012

Abstract

The Washington State Department of Natural Resources, Division of Geology and Earth Resources, houses many geologic and geophysical data collections, 16 of which have been inventoried in the National Digital Catalog. For our 2011 project for the National Geological and Geophysical Data Preservation Program (NGGDPP), we focused on our metal mine files, spreadsheets and databases.

Our objective was to create site specific metadata for submission to the National Digital Catalog and consolidate existing databases and spreadsheets into a single downloadable geospatial relational database that would include hyperlinks to scanned documents, reports and maps that were previously only available by physically accessing the paper files.

To accomplish this, our staff scanned our metal mine files and flat map collection covering 4,050 mining sites. We entered data for 29,436 scanned pages, compiled databases and spreadsheets into a single database which contains 106,246 hyperlinks to scanned documents and databases. This information was compiled into XML format in accordance with the National Digital Catalog specifications, and submitted for inclusion in the catalog.

Table of Contents

Abstract.....	1
Introduction	2
Project Goals	3
Methodology.....	4
Results.....	5

Introduction

The Washington State Department of Natural Resources, Division of Geology and Earth Resources (DGER), identified 16 collections of geologic and geophysical data which have been inventoried in the National Digital Catalog.

For our first year project (2009), we identified three collections that contain a wealth of high-value geologic and geophysical information, and were at imminent risk of degradation or loss, either through physical deterioration, lack of documentation, or disposal. These collections were rock core, rock cuttings and geotechnical reports. For year 2009 of the National Geological and Geophysical Data Preservation Program (NGGDPP), we produced XML metadata records for 47 boreholes that produced rock cores, 309 boreholes that produced rock cuttings, and 56,302 boreholes represented in 10,000 geotechnical reports for a total of 56,658 total metadata records submitted to the National Digital Catalog.

In 2010 our project continued the work began in 2009 on the geotechnical report collection, providing an additional 28,196 records to the National Digital Catalog covering 8,053 reports. The Division of Geology and Earth Resources also launched the Washington State Geologic Information Portal, an interactive mapping site designed to address one of the core requirements of the Geologic Data Preservation Plan of Washington, which is to make data available and accessible.

When the Washington Geological Survey was first established, one of the main directives was to collect drawings, maps, reports, minerals, and other information relating to the mineral industry and make this information accessible to the public. We continue this tradition by selecting the metal mine files for our 2011 data preservation project.

The information contained in these mining files is invaluable. Since 1890 the Washington Geological Survey has amassed an extensive collection of historical metal mining-related data on over 4,000 sites that include such information as ownership, mine maps, drilling records and assays, geochemical surveys, and production records. These files are unique, and are often the only source of information about mine workings, production, smelting processes, ownership, chemicals use, and the type of ore mined.

The importance of this information is underscored by the diversity of ways in which it is used, and the people who are using it. Not only is this information currently being used by the mining industry in one of the most extensive metal exploration programs that Washington has seen in years, it is also being used in litigation over metal mine contamination in the upper Columbia River and other areas of the state. Many others use the metal mining collection. It is routinely used by local, state, and federal agencies, lawyers, small miners, mining companies, researchers, mining and engineering students, realtors, educators, and the general public.

Project Goals

The purpose of this project was to preserve physical data by converting paper files to digital format, and consolidate these data with existing metal mine-themed spreadsheets and databases. These data would be consolidated into a relational database and FGDC compliant metadata would be written for the database. XML metadata would then be generated and submitted to the National Digital Catalogue.

At the beginning of the project DGER estimated that the number of items for which we would scan and create metadata for would be as follows:

- Approximately 40,000 pages of historical mining documents related to approximately 3,800 mining property sites
- Approximately 400 one-of-a-kind oversized metal mining plates, including cross sections, geologic maps, historical production information, drill-hole information, assays, and maps of mine workings of the largest mines in the state.

The project continued over a one year period beginning July 1, 2011 and ending June 30, 2012. It was anticipated that the scanning, digital enhancement of the images and data entry for the oversized flat files would be completed within the first three months of the project.

We anticipated the next seven months would be dedicated to scanning and entering data for DGER's historical metal mine files. Simultaneously we would build a schema for a relational geodatabase, and normalize data in existing spreadsheets and databases in preparation for consolidation into this single geodatabase.

The last 6 weeks we reserved for the data consolidation, writing complete FGDC metadata for the featureclass and all of the related tables in the geodatabase. Also during this time we would compile and submit the XML metadata to the National Digital Catalogue.

Methodology

We began our project by creating data entry spreadsheets for the oversized flat files. Double-click automation was used to facilitate the association of tabular data from known mine sites with the scanned oversized mining plates. We utilized drop-down menus and pick lists to normalize data as it was being entered into the spreadsheet. This cut down on the amount of data processing necessary later, during the quality control and quality assurance process.

After the oversized flat files were scanned in our wide format, continuous feed scanner, the images were processed in Adobe Photoshop. They were cropped and rotated if necessary, and other enhancements were made to the images to increase the visibility of the data in the image for the viewer.

The next step was to scan the historical metal mine files. First, we set up a similar spreadsheet as we did for the oversized flat files to use for data entry, and began scanning the files. It became apparent as we began this phase of the project, that we had underestimated the amount of time it would take to scan these files. This was due to there being many reports that were bound and we had not taken into account the time needed to unbind and rebind them.

We also found that there was no unique way to identify these mines in our files. Unlike our oil and gas exploration files, which have unique numerical identifiers, our mining files were filed by mine name. Some mine names are very popular. Names like “Eureka”, “Last Chance” and “Lucky Strike” have been used many times for different mine sites over the years. Also, file contents with the same name had in some cases been mixed together.

To assist us with these challenges we solicited help from Fritz Wolff, and Ian Hubert. Fritz Wolff has worked in the mining business for 10 years, and has worked for the Washington State Geologic Survey since 2001 as principal investigator for inactive and abandoned mine lands. Fritz’s first-hand knowledge of these mines was invaluable in refining locations, matching tabular data with mines and sorting out mine files. Ian Hubert is a geologist with good computer skills and helped with scanning and some GIS and data processing. This assistance was provided at no additional cost to the NGGDPP.

Coordinate information was ranked for accuracy according to the data source’s methodology for determining the coordinate, and the general accuracy of the coordinates in the source data. When multiple coordinates from different sources were provided for a single site, the location source with the best methodology and highest accuracy was used. Fritz checked the location accuracy on the highest profile sites using a combination of a program called Terrain Navigator, and first-hand experience.

Additionally, it became apparent during the scanning process that portions of some publications had been copied, and these duplicate copies were scattered throughout the files. Rather than scan the same document multiple times, we opted to scan the entire publication one time, and provide a link with page numbers where information for that particular mine could be found. This would reduce the amount of pages we estimated would be scanned, but save time scanning the duplicate pages, save storage space, and created a better product for the user.

During this same period of time a data base schema was developed for the metal mine geodatabase. Data in metal mine spreadsheets and databases were normalized in preparation for import and consolidation into the geodatabase.

In the last six weeks we consolidated the data into a single relational geodatabase that consists of a point featureclass and four related tables with corresponding relationship classes. Full FGDC metadata was written for all five data elements in the geodatabase. We moved the scanned files to a designated storage space used to host files where the public can access them, and created hyperlinks to the scanned documents. We performed QA/QC on the data, and generated XML metadata suitable for inclusion in the National Digital Catalog. Metadata was submitted to Rick Brown on June 29, 2012.

Two downloadable databases packages were to be placed on our website for download, the metal mine geodatabase, and the Inactive and Abandoned Mine Lands (IAML) database. The IAML database package was too large (240 mb) and needed to be reduced. The zip file contained many metal mine related photos of adits, mill sites, ponds, tailing piles etc. We removed the photos placed them in the same storage place area as the metal mine scans and created new links in the database. By hosting the photos on our image server we reduced the size of the download package to 10 mb. Updated links were forwarded to the NGGDPP.

Results

We submitted XML metadata for 4,050 sites that included links to 32,823 pages of scanned documents for the grant period July 1, 2011 to June 30, 2012. The delivery date of the metadata to the contracting official was June 29, which satisfied our contract requirements. The metal mine database was available for download within 14 days which also satisfies our contract requirements. The IAML data base was online within 30 days of the contract end, but was not listed as a deliverable in the 2011 grant proposal, so it did not fulfill any contractual obligations.

Below is a comparison of our project goals to our actual accomplishments during the grant period (July 1, 2011 to June 30, 2012).

Estimated:	Estimated number mine sites and documents to scan	~3,800 sites (~40,000 Pages)
Results:	Actual number of mine sites and documents scanned	4,050 sites (32,823 Pages)

We estimated that there were approximately 400 plates in the oversized flat files, but the actual amount was more than double what we had estimated – 848 plates. Our estimate of the number of pages to scan in the metal mine files was approximately 40,000 pages and we scanned 32,823 pages total. This was less than we had anticipated due to our decision to not scan the duplicate documents and publications repetitiously, and because the original number was an estimate. We had estimated that there were approximately 3,800 metal mine sites and we had 4,050 which is 250 more than we had anticipated.

In the future we plan to create a metal mine theme in the Washington State Geologic Portal, our interactive mapping application:

http://www.dnr.wa.gov/ResearchScience/Topics/GeosciencesData/Pages/geology_portal.aspx

The Washington State Geologic Information Portal utilizes the Arc Server interactive mapping application and is programmed with the Flex API, which is based on the Adobe Flash platform. One of the many strengths of Arc Server with Flex is serving large datasets to the user quickly. With this application the user can navigate through different themes; locate, overlay, and query data; create maps with titles, legends, and labels; create their own points and polygons; and export their maps as PDFs or geo-referenced TIFFs.